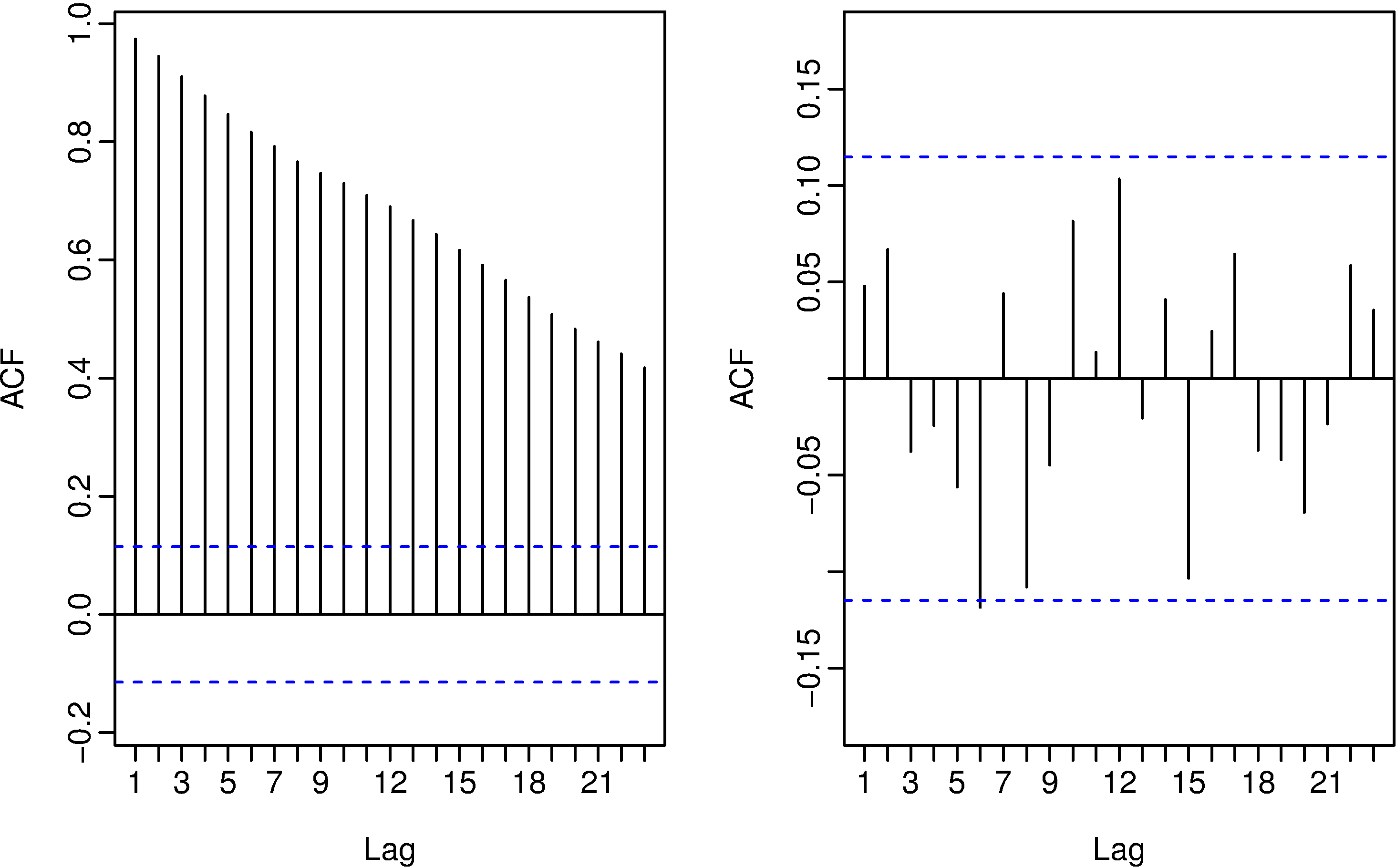
* The time series which we are dealing with is the Stationary time series as its properties doesn’t depend on the time at which the series is observed (the time series with trends/patterns and seasonality are Not-Stationary).
* White noise series is stationary- a time series with cyclic behaviour (but not trend or seasonality) is stationary. That is because the cycles are not of fixed length, so before we observe the series we cannot be sure where the peaks and troughs of the cycles will be.
* Mean shouldn’t be a function of time.
* Variance shouldn’t be a function of time but it must be constant
* Differencing – To make the non-stationary series stationary. The process involves computing the difference between the consecutive observations.
* Transformations such as logarithms can help to stabilize the variance of a time series.
* Differencing can help stabilize the mean of a time series by removing changes in the level of a time series, and so eliminating trend and seasonality.

ACF:  the ACF plot is also useful for identifying non-stationary time series. For a stationary time- series, the ACF will drop to zero relatively quickly, while the ACF of non-stationary data decreases slowly. Also, for non-stationary data, the value of r1 is often large and positive.



* The above are the ACF plots for non-stationary and stationary series respectively. First plot is decaying slowly that indicates it is non-stationary series and the second one shows exponential decay.
* Autocorrelation: Used to determine seasonality and stationary. Stationary series have a constant value over time.
* Moving average is purely random, so for a lag of either 1 or lag of 20, correlation should be constant or at least minimum.
* Stationary – ACF will drop to zero quickly
* Non-Stationary – ACF will decrease slowly, the value of r1 is large and positive
* Seasonal –
* Non-Seasonal- trend/irregular component
* If an additive model is not appropriate for describing this time series, since the size of the seasonal fluctuations and random fluctuations seem to increase with the level of the time series. Thus, we may need to transform the time series in order to get a transformed time series that can be described using an additive model. For example, we can transform the time series by calculating the natural log of the original data:

DIFFERENCING:

* Random Walk-
* Second Order Differencing - the differenced data will not appear stationary and it may be necessary to difference the data a second time to obtain a stationary series.
* Seasonal Differencing- A seasonal difference is the difference between an observation and the corresponding observation from the previous year.
* A moving average model is used for forecasting future values while moving average smoothing is used for estimating the trend-cycle of past values.
* ARIMA is applicable for NON-STATIONARY series, it reduces the NON-Stationary series to Stationary one.

|  |  |
| --- | --- |
| White noise | ARIMA(0,0,0) |
| Random walk | ARIMA(0,1,0) with no constant |
| Random walk with drift | ARIMA(0,1,0) with a constant |
| Auto-regression | ARIMA(*p*,0,0) |
| Moving average | ARIMA(0,0,*q*) |

**\*\* DETERMINATION OF AUTO-REGRESSIVE (AR)/MOVING AVERAGE (MA) MODEL:**

|  |  |  |
| --- | --- | --- |
|  | **ACF** | **PACF** |
| **AR** | Geometric | Significant till ‘P’ lags |
| **MA** | Significant till ‘P’ lags | Geometric |
| **ARMA** | Geometric | Geometric |
| **Non-Stationary Time-Series (Random-Walk process)** | Constant | Significant till ‘P’ lags |

* P, Q, D: Determine D(the order of Integration) , using adf (Augmented Dickey-Fuller)unit root test, P & Q using acf and pacf.
* If the series appears to be stationary around a constant, reject the null hypothesis. Then D=0 because the data series doesn’t need to be differenced to be made stationary.
* In adf test, the maximum lag order is the cube root of the sample size.
* ADF: The null-hypothesis for an ADF test is that the data are non-stationary. So, large p-values are indicative of non-stationarity, and small p-values suggest stationarity. Using the usual 5% threshold, differencing is required if the p-value is greater than 0.05.
* KPSS: Another popular unit root test is the Kwiatkowski-Phillips-Schmidt-Shin (KPSS) test. This reverses the hypotheses, so the null-hypothesis is that the data are stationary. In this case, small p-values (e.g., less than 0.05) suggest that differencing is required.
* Differencing:  
  Based on the unit test results we identify whether the data is stationary or not. If the data is stationary then we choose optimal ARIMA models and forecasts the future intervals. If the data is non- stationary, then we use Differencing – computing the differences between consecutive observations. Use ndiffs(), diff() functions to find the number of times differencing needed for the data &  to difference the data respectively.
* autocorrelations which measure the relationship between yt and yt−k for different values of k. Now if yt and yt−1 are correlated, then yt−1 and yt−2 must also be correlated. But then yt and yt−2 might be correlated, simply because they are both connected to yt−1, rather than because of any new information contained in yt−2 that could be used in forecasting yt.
* To overcome this problem, we can use **partial autocorrelations**. These measure the {relationship} between yt and yt−k after removing the effects of other time lags -- 1, 2, 3…, k−11, 2, 3…, k−1. So, the first partial autocorrelation is identical to the first autocorrelation, because there is nothing between them to remove.
* For ARIMA models, MLE (maximum likelihood estimation) is very similar to the *least squares* estimates.
* Value of the *log likelihood* of the data; that is, the logarithm of the probability of the observed data coming from the estimated model.
* Akaike’s Information Criterion (AIC), which was useful in selecting predictors for regression, is also useful for determining the order of an ARIMA model.
* Good models are obtained by minimizing either the AIC, AICc or BIC. Our preference is to use the AICc.
* The null-hypothesis for an ADF test is that the data are non-stationary. So large p-values are indicative of non-stationarity, and small p-values suggest stationarity. Using the usual 5% threshold, differencing is required if the p-value is greater than 0.05.
* *Kwiatkowski-Phillips-Schmidt-Shin (KPSS) test*. This reverses the hypotheses, so the null-hypothesis is that the data are stationary. In this case, small p-values (e.g., less than 0.05) suggest that differencing is required.
* ndiffs() which uses these tests to determine the appropriate number of first differences required for a non-seasonal time series.
* The order of the moving average determines the smoothness of the trend-cycle estimate. In general, a larger order means a smoother curve.
* Simple moving averages such as these are usually of odd order (e.g., 3, 5, 7, etc.) This is so they are symmetric: in a moving average of order m=2k+1, there are k earlier observations, k later observations and the middle observation that are averaged. But if m was even, it would no longer be symmetric.
* Combinations of moving averages result in weighted moving averages.
* STL is an acronym for “Seasonal and Trend decomposition using Loess”, while Loess is a method for estimating nonlinear relationships.